MARKOV PROCESSES AND RELATED PROBLEMS OF ANALYSIS[1]

1. The intimate connection between Markov processes and problems in analysis has been apparent ever since the theory of the former began to develop. It is not without reason that A. N. Kolmogorov's paper [39] (Russian translation [38]) of 1931, which is of fundamental importance in this domain, was entitled "On analytical methods in probability theory". The investigation of these connections also forms, to a large extent, the subject matter of A. Ya. Khinchin's book of 1933 on "Asymptotic laws of the theory of probability" [52] (Russian translation [51]).

In the fifties, and more particularly during the last five years, the theory of Markov processes entered a new period of intense growth. If previously the connections between probability theory and analysis were somewhat one-sided, probability theory applying results and methods of analysis, now the opposite tendency increasingly asserts itself, and probabilistic methods are applied to the solution of problems of analysis. Methods belonging to the theory of probability not only suggest a heuristic approach, but also, in many cases, yield rigourous proofs of analytic results. Applications of the methods of the theory

1 This paper is an expanded version of a survey read by the author at a meeting of the Moscow Mathematical Society held on October 20th, 1959, and devoted to the activities of the seminar directed by E. B. Dynkin at the University of Moscow.

of semigroups of linear operators have led to far-reaching advances in the classification of wide classes of Markov processes. New and deep connections between the theory of Markov processes and potential theory have been discovered. The foundations of the theory have been critically re-examined; the new concept of a strongly Markovian process has acquired a crucial importance in the whole theory of Markov processes. Intensive research on these lines is being done in the whole world, attracting many distinguished mathematicians: Feller, Doob, Hunt, Ray, Chung, Kac, and many others in the U.S.A.; Ito, Yosida, Maruyama, and their disciples in Japan; Kendall, Reuter, and others in England; Fortet in France. Soviet mathematicians are also taking an active part in this creative competition. A group of mathematicians cultivating the new approach in the theory of Markov processes is gathered in the seminar led by the present author at the University of Moscow. This survey covers the most important results obtained by members of the seminar from the academic year 1955–56 onwards, particular attention being paid to the latest results obtained during the last couple of years. Naturally, the work done by foreign mathematicians on the subjects studied by the seminar will also be discussed; however, in this respect, the survey cannot claim to be complete.

A short account of the indispensable general concepts of the theory of Markov processes is given in §1. A survey of the main lines of investigation followed by the seminar is given in §2–7. The concluding §8 contains information about the membership of the seminar and its history.

§1. Introduction

2. What is a Markov process? We shall begin with an important class of Markov processes, called diffusion processes, which describe a physical phenomenon known as *Brownian motion*. It is well known that particles of dye-stuff immersed in a liquid move chaotically, changing the direction of their motion all the time. This movement is due to collisions of the particles with molecules of the liquid. The first mathematical theory of Brownian motion was created by Einstein and Smoluchowski. In a contemporary form, due to A. N. Kolmogorov, this theory is shaped as follows: the main mathematical entity is a function $P(t, x, \Gamma)$, which represents the probability of a particle being in the set Γ after a length of time t from a moment when it was at the point x. Concerning the properties of this function, some assumptions are made from which it follows that

$$P(t, x, \Gamma) = \int_\Gamma p(t, x, y)\, dy,$$

where $p(t, x, y)$ is the fundamental solution of the parabolic equation

$$\frac{\partial p}{\partial t} = \sum a_{ij}(x) \frac{\partial^2 p}{\partial x_i \partial x_j} + \sum b_i(x) \frac{\partial p}{\partial x_i}. \tag{1}$$

This result makes it possible to apply the theory of differential equations to the solution of a variety of important problems on Brownian motion. On the other hand, many other, not less important, questions do not fit into this theory. For instance, one may want to know how quickly particles of dyestuff will be deposited on an absorbing screen; but it is impossible to solve this problem by means of the function $P(t, x, \Gamma)$ without making additional assumptions.

3. A more complete mathematical model of Brownian motion should give an account not only of probabilities involving one moment of time, but also of those which involve the whole course of the process; the whole trajectory x_t should be the object of the theory. The random character of the motion is expressed mathematically by the assumption that $x_t = x_t(\omega)$, where ω is an element of the set Ω, which is "the space of elementary events" on which a system of probabilistic measures P_x is given. The sets A on which $P_x(A)$ is defined are described as events associated with the process, and the value of $P_x(A)$ is interpreted as the probability of the event A under the condition that the motion began at the point x. In particular, one of the events associated with the process is $\{x_t \in \Gamma\}$. The probability $P_x\{x_t \in \Gamma\} = P(t, x, \Gamma)$ is called the transition function of the process; whereas in the first model it was regarded as the only mathematical characteristic of the model, it now occupies a subordinate position.

The mathematical entity we have arrived at is precisely a Markov process in the modern sense. It is a pair (x_t, P_x), where $x_t = x_t(\omega)$ is a function of $t \geqslant 0$ and of $\omega \in \Omega$, and P_x is a system of probability measures in the space Ω. The phase space to which the values of x_t belong is, in the case of Brownian motion, a domain of the three-dimensional space. In general, however, it is an arbitrary set E for which a system of "measurable subsets" has been defined. One essential condition has to be satisfied by the function x_t and the measure P_x: it is the Markovian principle that the future should be independent of the past when the present is known. More precisely, given the value of x_t, prospects of the future motion of the particle should not depend on its movement before the instant t.

4. The general scheme connecting Markov processes with analysis is based on the concept of a shift of a function defined over the phase space. The value of t being arbitrarily fixed, let $f(x)$ be a measurable function over the phase space. Then $f(x_t)$ is well defined over Ω; the integral of this function with respect to the measure P_x is precisely the value of the shifted function at the point x. Thus

$$T_t f(x) = M_x f(x_t) = \int_E P(t, x, dy) f(y),$$

where $P(t, x, \Gamma)$ is the transition function. The shift of a function is a linear operator. The Markovian principle implies $T_s T_t = T_{s+t} (s, t \geqslant 0)$, so that the operators T_t form a semigroup. Consider now the "operator of an infinitely small shift".

$$Af(x) = \lim_{t \to 0} \frac{T_t f(x) - f(x)}{t}. \tag{2}$$

This operator is called the *infinitesimal* operator of the Markov process. If Af is defined, $T_t f$ is the solution of the equation

$$\frac{\partial u}{\partial t} = Au \tag{3}$$

which satisfies the initial condition $u(0, x) = f(x)$.

For the diffusion process (which describes Brownian motion), the infinitesimal operator is given by[1]

$$Af = \sum a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum b_i(x) \frac{\partial f}{\partial x_i} \tag{4}$$

(for every x, the numbers $a_{ij}(x)$ form a positive semi-definite matrix). In this case, (3) is essentially equivalent to (1).

In general, the transition function of a process can be regarded as the fundamental solution of equation (3). But having at our disposal a system of measures P_x allows us to create, within the framework of our theory, a much greater variety of constructions and transformations than that which would be possible within the pure theory of differential equations. For instance, in the formula defining the shift operator T_t, one can replace the constant t by a random time τ. The operators T_τ can no longer be expressed in terms of the transition function. Furthermore, by means of such operators, one can express, for instance, the solution of Dirichlet's problem for an arbitrary elliptic equation in an arbitrary domain.

§2. General problems in the theory of Markov processes

5. These problems concern the foundations of the theory of Markov processes and have mostly a set-theoretical character. I shall briefly discuss three problems of this kind.

The first concerns the nature of the trajectories of the process. Consider, for instance, a diffusion process whose infinitesimal operator is defined by (4). Are all the trajectories of the process continuous? This question is of crucial importance not only in the study of Brownian motion, but also in the qualitative solution of any analytical problem connected with the differential operator (4). What is the answer to this question? In the first place, the question is not quite correctly stated. The point is that the set of trajectories is not uniquely defined by the infinitesimal operator or by the transition function. Therefore, a more

1 Strictly speaking, the operator A is given by (4) for all functions with continuous partial derivatives up the second order. However, its domain contains some less regular functions as well. Thus the infinitesimal operator is an extension of (4).

correct statement of the question would be this: is there a Markov process admitting an infinitesimal operator (4), and such that all its trajectories are continuous? The answer is "Yes". In the theory of Markov processes, a diffusion process is always understood to be one admitting an operator (4), and having the property that all its trajectories are continuous. Such a process with continuous trajectories is essentially defined by the operator A.

A general criterion allowing one to decide to which transition functions there correspond Markov processes with continuous trajectories was given by E. B. Dynkin in 1952 [35]. A little later it was found independently by Kinney [82]. The condition is very simple; however, it is only sufficient, and not necessary. In 1957, L. V. Seregin [43] deduced another, slightly more complicated, but stronger, criterion which, in a wide class of cases, is not only sufficient, but also necessary for the continuity of the trajectories.

On the basis of the Dynkin–Kinney criterion, a simple sufficient condition for the continuity of the trajectories can be given in terms of the infinitesimal operator. The essential part of this condition requires that the operator should have a local character, i.e. that $Af(x_0)$ should not vary when the function is modified outside a neighbourhood of x_0. Clearly, this condition is satisfied by all differential operators. Hence one can obtain the proposition previously mentioned about the continuity of the trajectories of diffusion processes. In papers by Dynkin, Kinney, and Seregin, beside conditions for the continuity of the trajectories of a process, conditions are also deduced for their continuity to the right, and for their having no discontinuities of the second kind. With respect to the trajectories of a special class of Markov processes, very subtle conditions for their continuity to the right were found by A. A. Yushkevich [57], [58].

6. A second problem arising in the theory of Markov processes concerns the domain of validity of the Markovian principle of the future being independent of the past when the present is known.

Decompose a trajectory of the process into two parts: up to the time τ when a set Γ is first reached, and after this time. Assume that x_τ is known. Is the knowledge of the trajectory before the time τ relevant to the prediction of the motion after the time τ? Physical intuition requires a negative answer. However, such an answer does not follow from the definition of a Markov process, since this definition involves a fixed time t, and not a random time τ. We describe as strongly Markovian those Markov processes for which the principle that the future should be independent of the past when the present is known applies not only to a fixed time, but also to a well-defined class of random times τ.

The first paper in which the strongly Markovian property of some processes was rigourously stated and proved was written by J. L. Doob [64]; it contained a discussion of a special class of Markov processes with denumerable phase spaces. More general processes with denumerable sets of states were investigated from this view point by A. A. Yushkevich [56] in 1953.

The study of strongly Markovian processes as a class in its own right was initiated in papers by E. B. Dynkin [22], [23], [26] and E. B. Dynkin and A. A. Yushkevich [36][1] in 1955–56. Dynkin showed that, starting from the strongly Markovian property, and imposing definite conditions of continuity on the trajectories of the processes, one can compute their infinitesimal operators. In their joint paper, Dynkin and Yushkevich were the first to give a general definition of a strongly Markovian process; they constructed examples of Markov processes which are not strongly Markovian, and deduced sufficient conditions for a Markov process to be strongly Markovian.

These conditions require that, in an appropriate topology, all the trajectories should be continuous to the right and that shift operators should transform continuous bounded functions into continuous functions. It is easily seen that diffusion processes satisfy both these conditions, and, therefore, are strongly Markovian.

The strongly Markovian property was further analyzed in a succession of papers (A. A. Yushkevich [57], [59], R. Blumenthal [60], E. B. Dynkin [27], G. Maruyama [84], P. Lévy [83], D. Ray [87]) which appeared during the last three years. In his very interesting paper, D. Ray showed that under quite general assumptions it is possible to extend the phase space of a Markov process in such a way as to make the process strongly Markovian.

The basic results concerning strongly Markovian processes, as well as the essential criteria for the continuity of trajectories, are discussed in E. B. Dynkin's monograph [33].

7. The third set-theoretical problem which I wish to mention concerns the introduction of an intrinsic topology in the phase space. Topology plays no part in the definition of a Markov process. The phase space E is an arbitrary abstract set in which a system of measurable subsets has been singled out. However, we have seen that the study of Markov processes requires the introduction of some kind of topology in the phase space. In 1959, E. B. Dynkin [31] proposed the following definition of an intrinsic topology: The set Γ is called open if, for every $x \in \Gamma$ a trajectory starting from x remains in Γ for a positive length of time with probability 1[2]. Interesting properties of the intrinsic topology have been proved for standard processes, a wide class of processes which includes all processes important for applications, in particular, all diffusions. For such processes it is shown that a point x belongs to the intrinsic closure of a set Γ if, and only if, a particle starting from x visits Γ with probability 1 during an arbitrarily short time interval. A function $f(x)$ is continuous in the intrinsic topology if, and only if, with probability 1, $f(x_t)$, regarded as a function of t, is continuous to the right; the last result is due to I. V. Girsanov [15].

1 The year 1956 also brought three American papers (Hunt [78], Chung [62], Ray [86]) investigating independently of Yushkevich and Dynkin, various forms of the strongly Markovian property for some special classes of Markov processes.

2 In the case of Brownian motion, which corresponds to Laplace's operator, this topology coincides with the topology that was previously investigated by H. Cartan [61] and J. L. Doob [65], [66].

I. V. Girsanov [15] and M. G. Shur [54] proved that the continuity of a function $f(x)$ in the intrinsic topology is invariant with respect to shifts. The condition that shifts should transform every continuous function into a continuous function plays a very important part in the theory of Markov processes; we have seen that it is one of a set of two conditions which are sufficient to ensure the strongly Markovian character of a Markov process. Roughly speaking, this condition means that trajectories starting from neighbouring points behave in a similar way. The part played by this condition was first pointed out by W. Feller, and this is why processes satisfying it are described as Fellerian. By proving that, in the intrinsic topology, every standard process is Fellerian, Shur and Girsanov obtained a result of fundamental interest.

The intrinsic topology is by no means the only interesting topology for the theory of Markov processes. In particular, I. V. Girsanov [15] proposed an interesting definition of a uniform structure connected with a process. For Brownian motion, this structure is induced by the usual Euclidean metric.

§3. The form of an infinitesimal operator. Generalized diffusion processes

8. Which operators are infinitesimal operators of Markov processes? This question is of crucial importance in the theory of Markov processes, because, under very general assumptions, the transition function can be built up from the infinitesimal operator in a unique way (see [25]), and the knowledge of the transition function allows one to obtain some insight into the whole class of processes corresponding to this function. On the other hand, the answer to the same question affects the analyst too, since it tells him which operators are susceptible of treatment by probabilistic methods.

An important tool for the investigation of the forms of differential operators is supplied by a general theorem due to E. B. Dynkin [22], [26]. Its statement is expressed by the formula

$$Af(x) = \lim_{U \downarrow x} \frac{T_{\tau_U} f(x) - f(x)}{M_x \tau_U} . \tag{5}$$

Here U is a neighbourhood of x, and τ_U the time of the first exit from U; the passage to the limit takes place when U is contracted into x.

One notices the analogy between this formula and formula (2), which defines the infinitesimal operator. However, despite the outward similarity of these two formulae, the passage from one of them to the other is far from trivial. In this passage, essential use is made of the fact that the process in question is strongly Markovian and continuous to the right. Strictly speaking, the basic theorem, as stated above, applies only to Fellerian processes, but in a slightly different form it can be extended to non-Fellerian processes as well.

The main term in the right-hand side of (5) can be expressed as follows:

$$T_{\tau_U} f(x) = M_x f[x(\tau_U)] = \int_E f(y) \Pi_x(dy),$$

where Π_x denotes the probability function of the point reached by the particle at the time of its exit from U. If the process is continuous, this probability is concentrated on the boundary of U. In this case, the recipe, given by (5), for obtaining the infinitesimal operator A is strongly reminiscent of the well-known recipe for obtaining the Laplace operator from the operator of averaging over a sphere; the only difference lies in the fact that in the general case the mean is taken with respect to a non-uniform measure, instead of the uniform measure used for the Laplace operator. It is natural to describe any operator obtainable by means of such a recipe as a *generalized second-order elliptic differential operator*. The adoption of this term is further justified by the fact that such operators have many of the properties of conventional elliptic operators, and also by the fact that if, in a domain, the limit in the right-hand side of (5) exists for functions giving coordinates and pairwise products of coordinates, then, for any twice differentiable function, this limit can be expressed by a conventional (possibly degenerate) elliptic differential operator.

Thus, if a Fellerian process is continuous, its infinitesimal operator is a generalized second-order elliptic differential operator. In this sense, *any continuous Fellerian process can be regarded as a generalized diffusion process*.

9. The contention that any Fellerian process with continuous trajectories is a generalized diffusion process is further strengthened in the case of processes on the straight line.

In this case the differential operator is found to be a generalized second derivative

$$Af(x) = D_v D_u \, f(x), \tag{6}$$

where $D_u f$ is the derivative with respect to the function u, i.e. the limit of the ratio of the increment of f to that of u. In this formula, u and v are arbitrary increasing functions; u must be continuous (v can be discontinuous). If u and v are twice differentiable, the operator (6) can be expressed in the form of

$$Af(x) = a(x) \frac{d^2 f}{dx^2} + b(x) \frac{df}{dx}, \tag{6'}$$

so that we are confronted with an ordinary diffusion process. It can be said that the general continuous one-dimensional process given by (6) is a diffusion process for which the coefficients $a(x)$ and $b(x)$ are (in a certain sense) generalized functions.

Formula (6) is an easy consequence of (5), but it was first obtained by W. Feller ([73], see also [77]) by an entirely different, purely analytical method. Feller's remarkable contribution provided one of the main stimuli for the development of the theory of Markov processes in the last few years.

10. Recently, a series of results was obtained concerning the properties of continuous processes corresponding to operators given by (6). In particular, A. D. Venttsel' [5] proved that the transition function of such a process can be expressed in the form

$$P(t, x, \Gamma) = \int_\Gamma p(t, x, y)\, dv(y),$$

where $p(t, x, y) = p(t, y, x)$ is the fundamental solution of the equation

$$\frac{\partial p}{\partial t} = D_v D_u p.$$

An important step forward was also made by A. D. Venttsel' [4] in connection with another question, which was investigated in an outstanding paper by I. G. Petrovskii [41] as early as the thirties. The probabilistic meaning of the problem is this: to find the order of magnitude of the greatest deviation, from the initial position x, of the moving particle during a time t with $t \to 0$. By means of a suitable transformation of the x axis, the general case can be reduced to that of $u(x) \equiv x$ (in which case $b(x) \equiv 0$). I. G. Petrovskii gave a complete solution of the problem for the process corresponding to the operator $\frac{d^2}{dx^2}$. It is easy to show that with any coefficient $a(x)$, behaving in a regular way, the overall picture will remain the same. A. D. Venttsel' investigated the general case of processes ruled by the operator (6), and showed that, here, various qualitative departures from Petrovskii's results were possible. Roughly speaking, in the case treated by Petrovskii, the greatest deviation of the particle from its initial position during a length of time t is of the order of $t^{1/2}$. Venttsel' showed that if the coefficient $a(x)$ has a singularity at the initial point of a motion, this deviation can be of any order t^α, where $0 < \alpha < 1$. For points of discontinuity of $v(x)$, this deviation can be of the order of t.

More general differential operators

$$a(x)\frac{d^2 f}{dx^2} + b(x)\frac{df}{dx} + c(x) f, \text{ where } c \leqslant 0.$$

can also be fitted into probabilistic schemes. To such operators there correspond Markov processes terminating at random instants. As recently shown by E. B. Dynkin [90], the general form of such terminating continuous processes on the straight line is obtained by replacing (6) by

$$Af = D_v D_u (qf), \tag{7}$$

where q is a function which is convex from above.

11. In the process of deducing (6) from the basic theorem (5), one obtains simple expressions for the functions u and v in terms of entities characterizing the process from a probabilistic view point, and having an intuitive meaning. If the motion takes place on a segment, these entities are: the probability $p(x)$ of attaining the right-hand end before the left-hand end when starting from x and the mean time $m(x)$ elapsing between the start from x and the arrival at the boundary.

Analogous entities with intuitive meanings can also be introduced for many-dimensional processes. In a survey read before the III All-Union Congress of Mathematicians, E. B. Dynkin [29] proposed the following problem: for which classes of many-dimensional processes do these two entities completely determine the process? In 1958, I. V. Girsanov [14] showed that one of these classes is formed by multi-dimensional diffusion processes with non-degenerate matrices $a_{ij}(x)$.

12. I have dwelt mostly on processes with continuous trajectories. However, quite a few substantial results have been obtained concerning processes with discontinuous trajectories. In particular, E.B. Dynkin [28] gave a full classification of jump processes. These are processes in which the whole motion proceeds in jumps, i.e. in which the particle remains in its initial position for some positive time, then jumps to a new point, then to another point, and so on. The main difficulty encountered in the study of such processes was due to the possibility of transfinite sequences of jumps.

§4. Harmonic, subharmonic, and superharmonic functions associated with a Markov process

13. I recall the general definition of a superharmonic function. A function $f(x)$ is *superharmonic* if

(a) the mean of f over any sphere centred at x is smaller than, or equal to, $f(x)$;

(b) $f(x)$ is continuous from above.

If the first condition is satisfied when the phase "smaller than, or equal to" is replaced by "equal to", f is described as *harmonic*; if it is satisfied when the same phrase is replaced by "bigger than, or equal to", f is called subharmonic. To fix ideas I shall confine myself to superharmonic functions.

To every Markov process one can attach a class of functions which are analogous to ordinary superharmonic functions. Conditions (a) and (b) are replaced by

(a′) $T_\tau f(x) \leqslant f(x)$ (τ being the first exit time from an arbitrary open set);

(b′) $T_{\tau_n} f(x) \to f(x)$ if $P_x \{ \tau_n \to 0 \} = 1$.

In the case of a diffusion process, corresponding to the Laplace operator, conditions (a′) and (b′) are equivalent to (a) and (b). The proof is far from simple, but it is easy to explain why (a′) implies (a). Let τ be the first exit time from the solid sphere bounded by S. In view of the invariance of the Laplace operator with respect to all rotations, a particle starting from the centre of the sphere will have a uniform probability distribution on the sphere S when it